

(Citation 3)

JP Patent Application Disclosure No.57-102699-June 25, 1982

Application No. 55-179945-December 18, 1980

Applicant: Matsushita Electric Industrial Co., Ltd, Tokyo (JP)

Title: VOICE RECOGNIZER

Detailed Description of the Invention:

.....

The switch (18) is connected to the side E at the registration time, and to the side R at the recognition time. The switch (19) controls as to whether the signal line (20) is transmitted to the switch (21), and is connected at the registration time. If constituted in this manner, when the voice word is generated twice at the registration time, the voice of the first time is stored in the registration pattern area (I) (17), and the voice of the second time is stored in the input pattern area (3), and the pattern matching unit (6) can calculate a degree of similarity with the voice of the first time and the voice of the second time. Although the judging unit (7) performs the same process as the judging unit of Fig. 4, at this time, the degree of similarity with the input pattern is calculated only with the registration pattern area (I), and hence, the Smax becomes the

degree of similarity with the registration pattern area (I) and the input pattern.

If the formula (1) is satisfied, the judging unit (7) allows the switch (21) to turn ON through the output line (20), and stores the content of the registration pattern area (I) (17) in the registration pattern area (II) (5) together with the category code of the input terminal (8). The registration pattern area (II) (5) is an area which stores the voice word to be recognized similarly to the registration pattern area of Fig. 4. If the formula (1) is not satisfied, the switch (21) is caused to turn ON by the output line (20). At the same time, whether the voice word is registered in the speaker by the output (9) is outputted. When not registered, the speaker further pronounces the same voice word. At this time, though it is the pronunciation of the third time, the degree of similarity with the pattern corresponding to the pronunciation of the second time and the pattern corresponding to the pronunciation of the third time is similarly calculated, and this is judged by the judging unit (7), and is subjected to the above described process. In this manner, until the degree of similarity becomes larger than a threshold value K, the same voice is pronounced as many times as necessary. Only when the degree of similarity becomes larger than the threshold value K, the registration pattern area (II) (5) is stored with the register

pattern, and after all, the registration pattern area is stored with a regular pattern, and an irregular pattern, for example, the pattern mixed with ambient noise is not registered.

.....

⑩ 日本国特許庁 (JP)

⑪ 特許出願公開

⑫ 公開特許公報 (A)

昭57—102699

⑮ Int. Cl.<sup>3</sup>  
G 10 L 1/00

識別記号

庁内整理番号  
7350—5D

⑬ 公開 昭和57年(1982)6月25日

発明の数 1  
審査請求 未請求

(全 6 頁)

⑭ 音声認識装置

⑯ 発明者 藪内秀和

門真市大字門真1006番地松下電  
器産業株式会社内

⑰ 特 願 昭55—179945

⑱ 出 願 昭55(1980)12月18日

⑲ 出 願 人 松下電器産業株式会社

⑳ 発 明 者 大賀英文

門真市大字門真1006番地

門真市大字門真1006番地松下電  
器産業株式会社内

㉑ 代 理 人 弁理士 大島一公

明 細 書

1. 発明の名称

音声認識装置

2. 特許請求の範囲

(1) 音声登録ボタンエリアと音声入力ボタンエリアとを設け、音声登録ボタンエリアにはあらかじめ認識すべき音声単語に対応する音声ボタンを格納しておいて、前記登録ボタンエリア内の音声ボタンと、前記音声入力ボタンエリア内の音声ボタンとのボタンマッチングを行なうことによつて入力ボタンエリア内の音声ボタンに対応する音声単語を認識する音声認識装置において、登録すべき音声は2回以上発声することにし、前記入力ボタンエリアに新たなボタンが入力される毎に、以前の入力ボタンエリア内のボタンを退避する登録ボタンエリア(I)を設け、入力ボタンエリア内の音声ボタンと登録ボタンエリア(I)内の音声ボタンとのボタンマッチングを行なつて類似度を算出し、その類似度が予め定められたシキイ値より大きい時、のみ前記登録ボタンエリア(I)内の音声ボタンあ

るいは前記入力ボタンエリア内の音声ボタンを登録ボタンエリア(II)へ転走することを特徴とする音声認識装置。

(2) 入力ボタンエリア内の音声ボタンと以前の入力ボタンエリア内のボタンを退避する登録ボタンエリア内の音声ボタンとのボタンマッチングを行ない、この時の類似度を $S_1$ とし、さらに前記入力ボタンエリア内の音声ボタンと、前記登録ボタンエリア内の音声ボタンとのボタンマッチングを行ない、最大の類似度を検出しこれを $S_2$ として、前記 $S_1$ がシキイ値より大でかつ $S_1$ が $S_2$ より大である時のみ、前記ボタンを退避する登録ボタンエリア内のボタンを登録ボタンとして登録ボタンエリアへ転送することにした特許請求の範囲第1項記載の音声認識装置。

3. 発明の詳細な説明

本発明は、あらかじめ登録された登録ボタン群と入力ボタン間とのボタンマッチングによつて、入力ボタンが登録ボタン群のどのカテゴリに最も似ているかを判断し、入力ボタンを識別するパタ

ンマッチング識別方法を音声に利用した音声認識装置に関するもので、その目的とするところは登録ボタンを登録する場合に正当でないものを排除するよう制御を加えようとするものである。

ボタンマッチング方式における音声認識方法について第4図に従つて説明する。発声は孤立音声単語であり、音声はマイク(1)によつて電気信号に変換され、特徴抽出部(2)によつて特徴が抽出されて特徴ボタンに変換される。又特徴抽出部(2)では音声区間の検出がなされ、発声された孤立音声単語に対応した特徴ボタンは入力ボタンエリア(3)に格納される。(5)は認識すべき音声単語の特徴ボタンとそのカテゴリコードを格納している登録ボタンエリアであり、スイッチ(4)がE(ENTRY)側と接続している場合には入力ボタンエリア(3)に格納されている特徴ボタンは入力端子(8)より入力されているカテゴリコードとともに登録ボタンエリア(5)へ転送される。あらかじめ、スイッチ(4)をR側にして、認識すべき音声単語を発声して、その特徴ボタンと、カテゴリコードを登録エリアへ格納す

(3)

なされてリジェクトされ、リジェクトに対応したコードが出力される。以上の処理は判定部(7)で行なわれる。

このような音声認識装置においては、登録に関して正当でない登録ボタンがそのまま登録されてしまうという問題があつた。登録すべき音声単語を発声した時に、たまたま雑音が加わつて来た場合、あるいはノドの調子がおかしい時に発声した場合等に得られた登録ボタンはかならずしも正当なものではない。正当でない登録ボタンのままで認識させれば、結果として誤認識が多発することとなる。

本発明は、登録されるボタンが正当なものか否かを判断し、正当なものである場合には登録し、正当なものでない場合には登録しないようにすることにより、従来より問題であつた正当でない登録ボタンによる誤認識を減少させたものである。

以下、実施例として示した図面に従つて、その構成を説明する。

(1)のマイク、(2)の特徴抽出部、(3)の入力ボタン

(5)

る。(4)は入力ボタンエリア(3)の入力ボタンと登録ボタンエリア(5)内の複数の登録ボタンとのボタンマッチングを行ない、登録ボタン毎に入力ボタンとの類似度を出力線(4)を介して出力するボタンマッチング部である。この時には、スイッチ(4)はR(Recognition)側と接続される。(7)はボタンマッチング部(6)からの出力である各登録ボタン毎の類似度を出力線(4)で受けるとともにその時のカテゴリコードを出力線(4)より受けて、入力ボタンがどの登録ボタンと最も似ているかを判別する判定部である。登録ボタン毎の類似度の中から最も大きい(最も類似性のある)値 $S_{max}$ を選出する。この $S_{max}$ があらかじめ定められたシキイ値 $K$ より大なる時、

つまり、 $S_{max} > K$  ..... (1)式

の時に入力ボタンは $S_{max}$ を出力した登録ボタンであると識別されて、その登録ボタンのカテゴリコードが出力(4)される。 $S_{max}$ が(1)式を満足しない時は登録ボタンエリア内に格納されている登録ボタンとは異なつた入力ボタンが入力されたとき

(4)

エリア、(5)の登録ボタンエリアおよび(6)のボタンマッチング部については第4図と同様な構成で、同じ動作である。

登録時には、同じ音声単語を2回以上発声することとし、カウンタ(4)は何回同一のカテゴリコードを有する音声単語が発声されたかをカウントするものである。特徴抽出部において、音声区間の検出はなされるが、音声単語の終端が検出されると出力線(4)を介してクロックを出力し、カウンタ(4)はそのクロックによつてカウントアップする。すなわち、音声単語を発声するたびにカウントアップされる。一方カウンタ(4)はリセット時検出部(4)からの信号によつてリセットされる。リセット時検出部(4)は入力端子(8)より入力されるカテゴリコード(登録されるべき音声単語に対応している)を受けて、そのカテゴリコードが変化したか否かを検出し、カテゴリコードが変化した場合にはカウンタ(4)をリセットする。同じ音声単語であるなら当然そのカテゴリコードも同一であるため、このような構成によつて登録する音声単語が何回発

(6)

声されたかを知ることが出来る。

カウンタ時の出力時はスイッチ時を制御し、登録すべき音声単語の発声が第1回目であるなら、スイッチ時をオフし、2回目以降ならスイッチ時はオンさせるように制御する。入力端子(8)のカテゴリコードを変化させない限りスイッチ時はオン状態となるため、認識時にカテゴリコードを変化させないようにすればスイッチ時はオン状態のままになる。認識時にはスイッチ時をオン状態にするように出力線(4)を制御しても良い。

登録ボタンエリア(I)の(1)の入力は、入力ボタンエリア(3)の出力と接続され、特徴抽出部(2)からの音声ボタンが入力ボタンエリア(3)へ格納される時に、入力ボタンエリア(3)に格納されていた以前のボタンは登録ボタンエリア(I)の(1)に転送される。従つて、登録ボタンエリア(I)の(1)の内容は、1つ前に発音された音声単語の特徴ボタンを格納していることとなる。

スイッチ時は登録時にはE側に、認識時にはR側に接続される。スイッチ時は信号線(4)をスイツ

(7)

(9)によつて話者に登録されたか否かを出力する。登録されない場合には、さらに話者は同一の音声単語を発音する。この時には8回目の発声となるが2回目の発声に対応したボタンと8回目の発声に対応したボタンとの類似度が同様算出され、判定部(7)で判定されて、前述したような処理がなされる。このようにすることによつて、類似度がシキイ値Kより大となるまで、何回でも同一の音声が発声することとなる。シキイ値Kより大になった時にはじめて登録ボタンエリア(II)の(5)には登録ボタンが格納されることとなり、結局登録ボタンエリアには正当なボタンが格納され、正当でないボタン、例えば周囲雑音が混在したようなボタンは登録されないことになる。

認識時には、スイッチ時をR側に接続しておき、かつスイッチ時をオフしておいて、スイッチ時は常にオフするようにしておけば良い。登録ボタンエリア(II)の(5)と入力ボタンエリア(3)とのボタンマッチングをボタンマッチング部(6)で行なつて、判定部(7)では前述したと同様な処理を行なえば、入力

(9)

チ(4)へ伝達するか否かを制御するもので、登録時には接続している。このような構成にしておけば登録時に音声単語を2回発生すれば、1回目の音声は登録ボタンエリア(I)の(1)へ格納され、2回目の音声は入力ボタンエリア(3)へ格納されることとなり、ボタンマッチング部(6)は第1回目の音声と、第2回目の音声との類似度を算出することが出来る。判定部(7)は、第4図の判定部と同様な処理を行なうが、この時には登録ボタンエリア(I)のみとしか入力ボタンとの類似度は算出されないため、前記した $S_{max}$ は、登録ボタンエリア(I)と入力ボタンとの類似度となる。

(1)式を満足しておれば、判定部(7)は出力線(4)を介してスイッチ(4)をオンさせ、入力端子(8)のカテゴリコードとともに、登録ボタンエリア(I)の(1)の内容を登録ボタンエリア(II)の(5)へ格納する。登録ボタンエリア(II)の(5)は、第4図の登録ボタンエリアと同様認識されるべき音声単語を格納しているエリアである。もし(1)式を満足していなければ出力線(4)によつてスイッチ(4)をオンさせる。又同時に出力

(8)

パターンの識別が可能となる。(4)は登録ボタンエリア(II)からの出力線である。登録ボタンエリア(II)の(5)には正当なボタンしか格納されていないため、従来に比べさらに認識率は良くなり、誤認識も減少することになった。

第2図は、判定部をさらに改良したものである。動作としては判定部(4)と判定部(7)は同様であるが、認識時と登録時でシキイ値Kを変化させる。スイッチ時は認識時にはR側に、登録時にはE側に接続され、認識時にはシキイ値は $K_R$ に登録時には $K_E$ になる。

$$K_E > K_R$$

としておくことにより、登録時にはさらにきびしく登録ボタンが正当か、否か判別されることになる。登録ボタンエリアには、より正当性のあるボタンしか登録されないことになる。このようにすればさらに認識率の向上、誤認識の減少が可能になる。

又、第8図に示すような構成も考えられる。入力ボタンマッチング部(6)は、入力ボタンエリア(3)

の内容と登録ボタンエリア(II)(5)の内容とは、認識時登録時に関係なくボタンマッチングを行なうようにし、登録ボタンエリア(I)(4)とは、登録時のみ入力ボタンとマッチングを計算するようにする。スイッチ部は従つて登録時のみオンし、認識時にはオフしている。さらに判定部(7)へは登録時のカテゴリコードを示している信号線部を登録時のみ加わる様な構成にしておく。スイッチ部は、登録時のみオンするようにしておく。カテゴリコード判定部は出力線部の内容と信号線部の内容を比較し、等しい時のみスイッチ部をオンするように制御線部を介してスイッチ部を制御する。

本発明はこのような構成をとつたので、登録時にはこれから登録しようとするボタンの他にすでに登録されているボタンともふくめて、ボタンマッチングを行なうことができ、従つて判定部(7)で選出される最も類似度の高い値  $S_{max}$  は、すでに登録されているボタンにおける類似度もふくめた値になり、この  $S_{max}$  を出力したカテゴリコードが出力線部に出力される。このコードがこれから

03

又、発声者は出力部の内容から登録されたか否かを判断することができる。指定したカテゴリコードが出力部に出力された場合には、登録されたことになり、逆に出力部のコードと指定したコードとが異なる場合には登録されないことを示す。出力部の内容が信号線部の内容と異なる場合には登録しようとする音声単語がすでに登録されている音声単語のどれと近いかを知ることができる。従つて、登録しようとする音声単語を変更する等の処理をすれば良い。例えばイチ(1)とシチ(7)ならシチをナナと発音するように変更する。

さらに判定部の処理を次に示すようにすれば、上記のような似かよつたボタン同士の登録を防ぐことができる。ボタンマッチング部から出力される類似度の中から2番目に高い類似度を  $S_{max2}$  とすると

$$S_{max} - S_{max2} > L \quad \cdots \cdots (2) \text{式}$$

という条件を加えれば良い。このようにすれば、すでに登録されているボタンとすくなくとも類似度が  $L$  以上のものしか登録されないことになる。

03

登録しようとするカテゴリコードであつた場合にはカテゴリコード判定部(6)によつてスイッチ部をオンし、登録ボタンエリア(I)(4)の内容を登録ボタンエリア(II)(5)へ転送するように出力線部に出力する。 $S_{max}$  を出力したカテゴリコードがこれから登録しようとするカテゴリコードでなかつた場合にはカテゴリコード判定部(6)によつてスイッチ部をオフするように出力線部に出力する。

又、最も高い類似度  $S_{max}$  が  $K$  より小の場合は、前述したと同様リジェクトに対応したコードが出力線部に出力され、当然これから登録しようとするカテゴリコードとは異なるためスイッチ部のオンはなされない。これから登録しようとするボタンが、すでに登録されているボタンの中に良く似ているボタンである場合(例えば、オカヤマとワカヤマ、イチ(1)とシチ(7)、ナハとナラ)にはすでに登録されているボタンの類似度の方がなんらかの原因(発音の不安定さ、あるいは周囲のノイズ)で、高いことがありうる。第8図のようにしておけばこのような場合には登録されないことになる。

04

従つて誤認識はさらに改良される。

又、第1図、第8図においては登録ボタンエリア(I)(4)を設けたがこれを設けず、登録ボタンエリアの内容の一部を消去する機能を設けても良い。登録時には登録エリアへ常に転送するようにしておき、登録時の判定の結果、登録してはいけない場合と判明した時に限つて、登録カテゴリコードに対応した登録ボタンの内容を消去するようにすれば第1図あるいは第8図と同様な処理が可能となる。又、第1図において、登録時に正当なボタンであると判断された時には、登録ボタンエリア(I)(4)内のボタンを登録ボタンエリア(II)(5)へ転送せず、入力ボタンエリア(3)の内容を登録ボタンエリア(II)(5)へ転送するようにしても良い。

本発明は上記のように雑音、あるいは発音の不安定さ等による原因で生じた正当でないボタンは登録されないために、誤認識を減少することができ、登録時に制限を加えることにより、似かよつたボタンの登録を防ぐことができ、安定した認識ができる効果がある。又、判定部のシキイ値を登

録時には認識時よりきびしくする方向にすれば、より正当なボタンしか登録されないため、誤認識がさらに減少する等の効果が得られる。

#### 4. 図面の簡単な説明

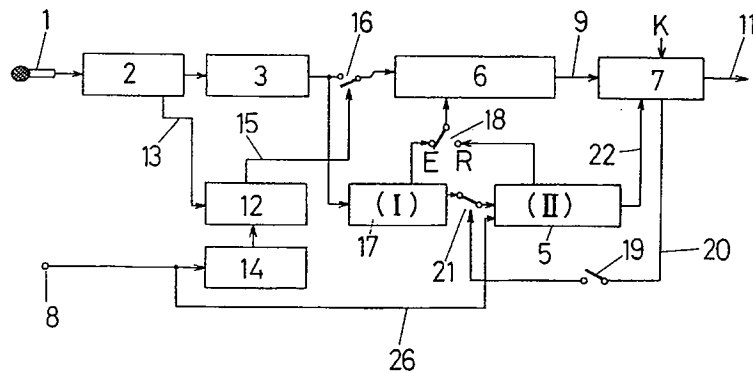
第1図は本発明装置の構成図、第2図は判定部の他実施例を示す構成図、第3図は本発明装置の他実施例を示す構成図、第4図はボタンマッチング方法による音声認識装置の構成図。

(3) …入力ボタンエリア、(5) …登録ボタンエリア  
(II)、(6) …ボタンマッチング部、(7) …判定部、(7) …登録ボタンエリア(I)

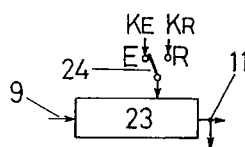
代理人 弁理士 大 島 一 公

05

第1図

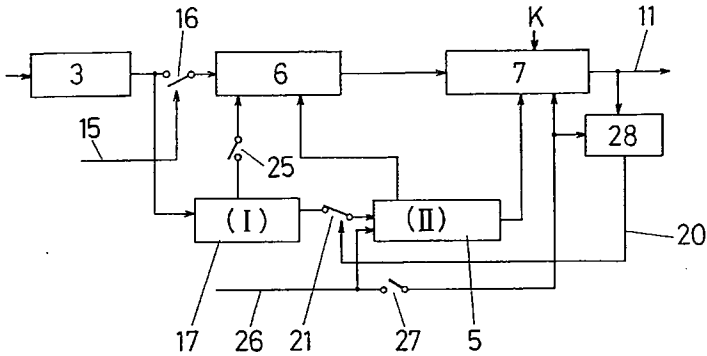


第2図





第 3 図



第 4 図

